

(TM)

Distribution rights by Intelligenetics, Inc.

- n.a. database search, using Smith-Waterman algorithm

ot generated

GCCGGTCCGGCGCGGAGGT..

TABLE default

Database 0; Que

362067 seqs, 549138275 bases x 2

Listing first 45 summaries

emb1-new3

genbank99

99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9

122:part1

Mean 11.533; Variance 4.446; scale 2.594

...ved by analysis of the total score distribution.

C	45	22	1.4	257912	103	PBU42580	Paramedium bursaria	C	2.15e+01
---	----	----	-----	--------	-----	----------	---------------------	---	----------

ALIGNMENTS

DEFINITION Sequence 5 from patent US 5569830.

NID 918190

SOURCE

ORGANISM

REFERENCE

TITLE

JOURNAL

FEATYPES

Source

BASE COUNT

100

F					
FT	Sequence	354 BP; 106 A; 42 C; 41 G; 54 T; 111 other;	HQAAKVPENLVX"		
D	Query Match	1.6%; Score 25; DB 8; Length 354;			
E0	Best Local Similarity	23.7%; Pred. No. 3,26e-01;			
	Matches	23; Conservative 38; Mismatches 36; Indels 0; Gaps 0;			
Dd	99	igcdagrgaryhyhdwargwdydgwcgykaayaagcwgawgmargaywasrttygtaar 158			
Cp	799	GGAAGGTCGTCTGTGTGACGTGTGTGGCGTTTATCCTCTACACACTCTGCCTTA 740			
Dd	159	gawdkrcgcgaacccwyrrrctmctcgagtmcddaar 195			
Cp	739	CTTGCTGCCAAGCAATTGTCTACCAGAACAATAAG 703			
RESULT LOCUS	7	S66477 1441 bp mRNA VRT 17-DEC-1993			
DEFINITION		glial fibrillary acidic protein [clone 2] [Cyprinus carpio-carp.			
SESSION		brain, mRNA Partial, 1441 nt].			
NID		S66477			
KEYWORDS		g435738			
SOURCE		.			
ORGANISM		common carp brain.			
		Cyprinus carpio			
		Eukaryotes; mitochondria; eukaryotes; Metazoa; Chordata;			
		Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
		Osteichthys; Cypriniformes; Cyprinidae; Cyprininae; Cyprinini;			
REFERENCE		1 (bases 1 to 1441)			
AUTHORS		Cohen,I., Shan,Y. and Schwartz,M.			
TITLE		Cloning and characteristics of fish glial fibrillary acidic			
JOURNAL		protein: implications for optic nerve regeneration			
MEDLINE		J. Comp. Neurol. 334 (3), 431-443 (1993)			
REMARK		93388923			
		GenBank staff at the National Library of Medicine created this			
		entry [NCBI gibbsq 139069] from the original journal article.			
		This sequence comes from Fig. 1.			
FEATURES		Location/Qualifiers			
source		1..1441			
CDS		/organism="Cyprinus carpio"			
		1..636			
		/partial			
		/note="Description: glial fibrillary acidic protein, GFAP,			
		Method: conceptual translation with partial peptide			
		sequencing. This sequence comes from Fig. 2"			
		/codon_start=1			
		/product="glial fibrillary acidic protein"			
		/db_xref="PID:g435738"			
		/translation="VDLDVSKPDDTLAKETRAQFEMATSMMOETEHWRSKPADITL			
		YDASIVARLEDIOMLKEMMARHLOEYDLWVKALDEIATYRKLLEGESRTLPV			
		ONFNLFDRDISLTDTKLPEAHVRSYVRVERDSEILNESTERNDLP"			
BASE COUNT		420 a 273 c 347 g 401 t			
ORIGIN					
		Query Match			
		Best Local Similarity 77.8%; Pred. No. 3,26e-01;			
		Matches 35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;			
Dd	438	gaagctgtcgaaggagaagaagaatcacacttccgtyca 482			
Cp	931	GAGCTGCTCGAAGGTGAGGTAGCATGCAATGTCCGCTGCA 887			
RESULT LOCUS	8	CRAGF1 1933 bp mRNA VRT 01-SEP-1993			
DEFINITION		Carassius auratus (GFAP-1) mRNA, complete cds.			
ACCESSION		L23876			
NID		g388622			
KEYWORDS					
SOURCE		Carassius auratus adult retina cDNA to mRNA.			

ORGANISM	Carrassius auratus
AUTHORS	Eukariotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Osteichthys; Cypriniformes; Cyprinidae; Cyprininae; Carrasius.
TITLE	1 (bases 1 to 1933)
JOURNAL	Glasgow,E and Schechter,N. Nucleotide sequence of a GFAP - like intermediate filament cDNA from goldfish retina unpublished (1993)
FEATURES	location/Qualifiers
source	1..1933
	/organism="Carrassius auratus"
	/dev_stage="adult"
	/tissue_type="retina"
	20..1099
	/gene="GFAP-1"
	/note="putative"
	/codon_start=1
	/db_xref="PID:g386623"
	/translation="MGINDRFASYIEKVERLEQONKMVAELNLGRKEPSRLDIOV EELSELROYDLGNAGKARLEIERDNLSADLTKLOJENALROAEKNILNTFRODD VDEALNRVOLEKIDALODEISFLRVEHEEMROLGOILAQOYHVDLVSKPLDTT AKETRAPOFMATSMNOETFEWRSFALTDAGNAENALPARKAENAYRPIOLG LTCSLESRGNSNESLERQLBMEBRRIETRTGYDIYARLEDEIOMIKREMARLOET ODLINVLALDIELIARYKLKGESSEITVPVNFTLNLPFRDISLDTRKPEAVHKRS IVRVETVRGEIKESIESTERKKDP"
BEGIN	561 a 395 c 507 g 470 t
BASE COUNT	
ORIGIN	
Query Match	1.6%; Score 25; DB 49; Length 1933;
Best Local Similarity	77.8%; Pred. No. 3.26e+01;
Matches	35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Db	901 gaagctgtcgaagagagaagacgatcattccgttcga 945
Cp	931 GAAGCTGCTCGAAGGTGAGTGATTGCATGTCCTCAATGTCGGCTGCA 887
RESULT	9
LOCUS	RMAP1B5 7095 bp RNA ROD 21-OCT-1992
DEFINITION	R.norvegicus mRNA for microtubule associated protein IB.
ACCESSION	X60370 X60371 X60550
NID	957618
KEYWORDS	MAB1B gene; microtubule-associated protein.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus
	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorphae; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 7095)
AUTHORS	zauner,W., Kratz,J., Staunton,J., Feick,P. and Wiche,G.
TITLE	Identification of two distinct microtubule binding domains on recombinant rat MAP 1B
JOURNAL	Eur. J. Cell Biol. 57 (1), 66-74 (1992)
MEDLINE	92347374
REFERENCE	2 (bases 1 to 7095)
AUTHORS	Wiche,G.
JOURNAL	Direct Submission
TITLE	Submitted (07-AUG-1991) G. Wiche, Inst of Biochemistry, University of Vienna, Waehringerstrasse 17, 1090 Vienna, AUSTRIA
REMARK	revised by (3)
REFERENCE	3 (bases 1 to 7095)
AUTHORS	Wiche,G.
JOURNAL	Direct Submission
TITLE	Submitted (07-AUG-1992) G. Wiche, Institute of Biochemistry and Molecular Biology, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, AUSTRIA
FEATURES	location/Qualifiers
source	1..7095
	/organism="Rattus norvegicus"
	/strain="Sprague-Dawley"
	/dev_stage="adult"

BASE COUNT	2124 a	1856 c	1799 g	1316 t	/tissue_type="Brain" /cell_type="C6 glioma" /db_xref="PID:9510544"
Query Match	1.6%;	Score 24;	DB 92;	Length 7095;	
Best Local Similarity	76.1%;	Pred. No. 1,38e+00;			
Matches	33;	Conservative	0;	Mismatches 11;	Indels 0;
Gaps	0;				
Db	1004	gcaaggaaatgcagattatcattcagtcagagtcgactggaacccaaca	1049		
QY	305	gcaaggagctgcagtagcgaacagagagtcgacatgcacccacaa	350		
RESULT	10				
LOCUS	PSCHSAB	7322 bp	DNA	PLN	06-JUL-1994
DEFINITION	P. sativum gene for chalcone synthase.				
ACCESSION	X80007				
VERSION	g510542				
ORGANISM	chalcone synthase; CHS-1A gene; CHS-1B gene; class 1 gene.				
REFERENCE	pea.				
AUTHORS	Pisum sativum				
JOURNAL	Eukaryotes; mitochondrial eukaryotes; Viridiplantae;				
REFERENCE	Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;				
AUTHORS	Magnoliopsids; Rutanae; Sapindales; Fabaceae; Papilionoideae;				
TITLE	Pisum.				
JOURNAL	1 (bases 1 to 7322)				
FEATURES	2 (bases 1 to 7322)				
SOURCE	Direct Submission				
	Submitted (01-JUL-1994) R. P. Hellens, John Innes Institute, Colney Lane, Norwich NR4 7UH, UK				
	Location/Qualifiers				
	1. 7322				
	/organism="Pisum sativum"				
	/strain="J1 813"				
	/clone_lib="labda GEM-11"				
	/clone="B"				
	/map="linkage group 3"				
	<1040..1217				
	/gene="CHS-1A"				
	/number=1				
	join(<1040..1217,1328..>2319)				
	/gene="CHS-1A"				
	join(1040..1217,1328..2319)				
	/gene="CHS-1A"				
	/EC_number="2.3.1.74"				
	/codon_start=1				
	/product="naringenin-chalcone synthase"				
	/db_xref="PID:9510543"				
	/translation="MTVNEIRAOARSPATVPFVIGTATPNCVCSQSYPPYPERITNSQKTEKKEKQKCDKSMIKRMYHLEILKENPSCIEFMAISLDRKQMYVEPKLKEATKAIKENGQPKSKITHLFCITSSVDMPGADYQJLKLGLRPYKRYMQQGCPAGGTVALAKDLAENKNGARVLVCSITAVTFRGPSDTHLDSLVGALFGDAAYIVGSDPLDVEKPLFEIYMTAQTIVPDEGAIDGHLRAGLTFHLKDPYSIVKNIKRALVEAPLPINISDYSNLFVIAHGGPAIILDOVEAKLGIKORKMQATRVHVEGMSASACVLTFLDEMRKSKEDGLATTBEGLEMGVLFGRGPETVETVYLHSMAT"				
	1218..1327				
	/gene="CHS-1A"				
	/number=1				
	1328..>2319				
	/gene="CHS-1A"				
	/number=2				
	join(<4336..4513,4642..>5633)				
	/gene="CHS-1B"				
	join(4336..4513,4642..5633)				
	/gene="CHS-1B"				
	/EC_number="2.3.1.74"				
	/codon_start=1				
	/product="naringenin-chalcone synthase"				
	/db_xref="PID:9510544"				

REFERENCE	AUTHORS	JOURNAL	TITLE	BASE COUNT	ORIGIN
exon				2350 a	1307 c 1264 g 2401 t
intron					
exon					
Query Match					
Best Local Similarity	1.6%	Score 24;	DB 64;		
Matches	31;	Conservative	0;	Mismatches	7;
				Indels	0;
				Gaps	0;
Db	654	aaagtgtaacctctcatgataaccacccaatggtc 631			
Qy	1404	AAAGGACTACATCTCTCCATATAAACCCAAAGGTT 1441			
RESULT	11				
LOCUS	CEC29F3	18632 bp	DNA	INV	31-JAN-1997
DEFINITION	Caenorhabditis elegans cosmid C29F3.				
ACCESSION	281043				
NID	g1627634				
KEYWORDS					
SOURCE	Caenorhabditis elegans.				
ORGANISM	Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda; Secernentea; Rhabditiia; Rhabditiida; Rhabditoidea; Rhabditiidae; Pelodierinae; Caenorhabditis.				
REFERENCE	1 (bases 1 to 18632)				
AUTHORS	Matthews, L.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jesusanger.ac.uk or rwenematode.wustl.edu				
REFERENCE	2 (bases 1 to 18632)				
AUTHORS	Wilson, R., Alnough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A., Saunders, D., Showkeen, R., Smalton, N., Smith, A., Sonhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterson, R., Watson, A., Weinstock, L., Wilkinson-Spoat, J. and Wohlman, P.				
TITLE	2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans				
JOURNAL	Nature 368 (6466), 32-38 (1994)				
MEDLINE	94150718				
COMMENT	Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is NOT necessarily the entire insert of clone C29F3. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. Coding sequences below are predicted from computer analysis, using the program GeneFINDER (P. Green, ms in preparation), and other				

FEATURES
SOURCE

available information.
Location/Qualifiers

1.18632
/organism="Caenorhabditis elegans"
/clone="C29f3"
/chromosome="v"
complement(join(158..344,421..598,647..698,743..796,
845..1128,1176..1414,1456..1544,1591..1788,2036..2201,
2721..2787))
/note="protein predicted using GeneFinder"
/codon_start=1
/db_xref="PID:e299269"
/product="C29f3.d"
/db_xref="PID:g1813905"
/translation="MTPADNRYVSNALAEFVPGHSSRRRRMETSLOGLARPD
KTIISFRHCHRVETWCGRIANRLEPTIQLTSSMTLLVSGLSYMWQVPSRCG
ESTIPSELEILPSGHPVLCARPTCFGHPKGYDLPPTAKSLNRLDLGLRDSLE
TYPEPDSKIKYQNSCEPQSCDSDQWVGLEPETDAFOVAYOCCTYAPL
RESIDRIATVAGEIYIGEVYQNSQDQVNSIEKSMDEGEVYEVNIRFAC
LDPNARRIDEVSESENTIRKVGOKPDAQANVAVNAPTEAGTDEGVDDGQTVY
IEETIAOGGFVENETVAPRACPGFOAGPERRGACQGPAPQPPPOQFPQNC
PVYQDPVQFPAQVPIAPAPAGMQLHCFPADAEVNYEKGVRKDELEGVQVQL
HGKETTSPVAKYWLHDPDEAEFEVLENGESFTLEKLVFATDQOQVVKNLDDL
NPSTGKINGCEFFMAQPEVRSILCY"
complement(join(5522..5692,5740..5834,5888..6062,
6642..6740))
/note="protein predicted using GeneFinder"
/codon_start=1
/product="C29f3.c"
/db_xref="PID:e299497"
/db_xref="PID:g1813904"
/translation="MTIILIFAFECMSVEGCIPTMPPEPVVYVYVPCPAGWFSQ
RANGIMCYIAKPPACWTSBOQYCDQNSQSPFNGESAERQFQIDALAAHLAS
QFLHGIIRSCAAPPCATDPYVQNGVSNQNFANDYIQLYDSSGGLSMDLAKGE
INDITCALIRVSCGYAA"
complement(join(7079..7249,7299..7387,7433..7601,
8197..8292))
/note="protein predicted using GeneFinder"
/codon_start=1
/product="C29f3.b"
/db_xref="PID:e299268"
/db_xref="PID:g1813903"
/translation="MMWRLVLFPSASOLSEGLPMYPPPEPVVYVPCPAGWFSQ
ATGIMCYIPATPGAGWTPOACQAGANNGESAERQFQIDALAAHLAS
HIGMRCARCTVNDPFWLVNGVNDNTFANDYISLDLGDCLSMGLGNQYNDIT
CDATAYSCGPPAA"
join(10724..10982,11039..11394,12940..13052,13183..13295,
13344..13893,13948..14262,14952..15141)
/note="protein predicted using GeneFinder"
/codon_start=1
/product="C29f3.g"
/db_xref="PID:e299270"
/db_xref="PID:g1813906"
/translation="MPEPSTFFEDHSRDELFQENVFEGFLVAFHDSVLAFTIS
LNRIISVMPLEFYTKFMNLKTYKLFIFAWMLISGLSILKSLNETKKSBSBAQOFS
ILASVCFEIRGVDLFRNSCIYIMMLIDVETLIKYLISNETKKSBSBAQOFS
RDRFLQATYQGSILFLATVVFPTPEVSNWTFPGNSCGRWIVPQYFATIL
TNTISSNNSAPVQAVDQWQVQVSTSESVSEVNGENIKLSGNRNVQGFSL
FMDYQFIOIKYLFNSATLPLYSNMPDQVITADTRISATIVPGVDSQYSLR
GVLEFDGTNRSTYGLTGLTGSKQVSYSTISYSDSNVANNPSPLADYEN
TKSICRFEWILCPDDQDGLISLDTGKSAALQMLNKKFSTSVLKTFGKGLAEV
DGVYTKNLITTVYVNSSENLPOEFGRLLTVLFETISALITTRDDADDETTSIG
RKGTISDSIGNSQFAFHINAPVKNPDIKVAIADTADSGSTLLVYCRITVELR
ELSENSTPLPLNTTFHGNFLHIEYIPNTDSKLFLENDIKSVTSRSEFFACV
IVFLFLF"

complement(join(15422..15713,16068..16591,16640..17140,
17376..18179))
/note="protein predicted using GeneFinder"

/codon_start=1
/product="C29f3.a"
/db_xref="PID:e299496"
/db_xref="PID:g1813902"
/translation="MNETLDVRVRLKRLFLFSYKQQLSDQSVKAIIVMSGKPNSEVAG

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

ADIDMEKERTAGVSNLRESGOKLITELISOKPIYAAIMSGMGGLEIATACHR
IANDKRTILGEPVETGIMPBGQGRTRKRTFVQVVDLTLTGKIKANKAKTET
VDRTIPLGIGICTSTETTKRYLEILVQSAERANKLITNRKGVHNQAVMTS
KFLVDNVLIMAKNRKLIKLTNGNPAKLLIDVAVRYLDPKNGVEAEKAFGLSQT
FOSKALFGEFGEINAKNRKYGKGLPYNEIAVYVAGMGAGIANVTINKGIRVILDA
NOAGVERGONHVAATHLNROKROKISLESEKIYNHLVPTIDYSAMNADVLEAVE
DLPLKHVIOIENRVENPTIISNTSALIKOIAAASSSDVIGWHSVPEKMO
LEITTHGTSKREITATAGIIOKOGKLVYVVKCCPGFFVRCISPMASEIVRLIOEY
EPSELDLTKTRFEPVGAATLADENAGIDVAEHVARYLGLRPHGGSVDLSELYR
AGHGRKTSKGTIEVYGDGASKRVNQEAKEKYLTPIKSVSPEDRQLRVSE
VNEALLICEGVISPSDGDIAVYFGLEFPFVGPPRFVLYGAGKLVAAAMRFAGV
YESVQFEPCCOLLRNHAKSGQFYS"

BASE COUNT 5872 a 3349 c 3316 g 6095 t
ORIGIN
Query Match 1.6%; Score 24; DB 35; Length 18632;
Best Local Similarity 90.0%; Pred. No. 1.38e+00;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 13078 ggaattaaagaactatgatgtggaataa 13107
Qy 1377 ggtactaaagaactatgatgtggaataa 1406

RESULT 12
LOCUS CER09D4 47745 bp DNA INV 06-NOV-1996
DEFINITION Caenorhabditis elegans cosmid T09D4.
ACCESSION Z81590
NID g1666006
KEYWORDS HTGS; PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdilia; Rhabdilita; Rhabdilitina; Rhabdilitoidea;
Rhabdilitae; Caenorhabditis.
1 (bases 1 to 47745)
AUTHORS Unknown.
REFERENCE Direct Submission
TITLE Submitted (04-NOV-1996) Nematode Sequencing Project, Sanger Centre,
JOURNAL Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
joesanger.ac.uk or rwenematode.wustl.edu
COMMENT Order of segments is not known; 800 nt's separate segments.
Cosmid-T09D4; Contig ID=00659; Length=42147; Status=Unfinished
Cosmid-T09D4; Contig ID=01673; Length=4019; Status=Unfinished
Cosmid-T09D4; Contig ID=01596; Length=1665; Status=Unfinished
Cosmid-T09D4; Contig ID=01600; Length=1271; Status=Unfinished.

FEATURES
source 1.47745
/organism="Caenorhabditis elegans"
/clone="T09D4"
/chromosome="v"
BASE COUNT 14131 a 8322 c 8318 g 14134 t 2840 others
ORIGIN
Query Match 1.6%; Score 24; DB 38; Length 47745;
Best Local Similarity 90.0%; Pred. No. 1.38e+00;
Matches 27; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 32078 ttcttcacatcagattcttcaattcc 32107
Cp 1406 ttcttcacatcagattcttcaattcc 1377

RESULT 13
LOCUS HS799F10 115419 bp DNA HTG 13-JAN-1997
DEFINITION Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone
799F10; HTGS phase 1.
ACCESSION Z82245
NID g1780984
KEYWORDS HTG; HTGS; PHASE1.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 115419)
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiinae; Homo

TITLE Direct Submission
 Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton,

COMMENT
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
IMPORTANT: This sequence is unfinished and does not necessarily

sequence: BK799F10 Contig_ID: 02214 Length: 19570 bp.

```

*** WARNING: Phase 1 High Throughput genome sequence ***
***
* This sequence is unfinished. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number
***

```

FEATURES
source

BASE COUNT	a	c	g	t	others
ORIGIN	24169	28425	26764	23990	12071

Query Match	1.68;	Score 24;	DB 34;	Length 115419;
Best Local Similarity	80.0%;	Pred. No. 1.38e+00;		
Matches	32;	Conservative	0;	Mismatches 8; Indels 0;

Db 109739 tctttgtttgttttttccactctgcagctgcaag 109778
 |||||
 |||||
 CP 833 TCTTTGTTTGATGTTTCATACTTCAGCAGCTGGAAG 794

RESULT	14
LOCUS	D90904 150894 bp DNA
DEFINITION	Synechocystis sp. PCC6803 complete genome, 6/27, 650555-781448

KEYWORDS formaldehyde dehydrogenase; ABC transporter; DNA ligase; DNA polymerase III alpha subunit; Gmms protein; Uvr transcriptional regulator; Mg protoporphyrin IX; N-acetylmethionine aminotransferase; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 5; acetate kinase; aspartate kinase (ASP); cation-transporting ATPase; cell division protein FtsH; chemotaxis protein CheA; cytochrome oxidase d subunit I; cytochrome oxidase d subunit II; dienelectone hydratase; dihydroflavonol 4-reductase; dihydrocytochrome pyrophosphorylase; esterase; ferrous iron transport protein B; fibrillin; formaldehyde dehydrogenase (glutathione); high light-inducible protein; leader peptidase I; lysostaphin; malonyl coenzyme A-acyl carrier protein transacylase; methyl-accepting chemotaxis protein (MCP); oligopeptide transport ATP-binding

MEDLINE

TITL

REFERENCE
AUTHORS

TITLE
JOURNAL

**SOURCE
ORGANIS**

REFERENCE
1 Eubacteria; Cyanobacteria; Chroococcales; Synechocystis.
1 (bases 1 to 150894)

COMMENT potential protein coding regions were assigned on the basis of similarity search of the ORFs and Genemark analysis.

tRNA

CDS
 complement(3235..4326)
 /note="ORF_ID:s110997"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652239"
 /transl_table=11
 /translation="MAPKNSCHNGDMKNSRLNGEMIVAPYQSFHIGLLGALASVM
 /PSACAIPEPPEAVNOQVNGDROEQVLAQNSDPTPTPEFDPSNDEKRLNGYKSG
 DVSASISYRGEKHGCESEFPTVTKFKKITYATLEGSSEPMIALLODITDMDRNPY
 PAULFATYRGAHCNCEYKVFETSNQSGMSVRDGFRTNGGHPADLNDKMDPEVA
 VDNRLILFSTAGSHAPQIWLQNGQVYDVSFEFSOFIRENNQSENKLPETIVA
 QDSEKNGFLAAVANKALIGLEDEGQVTLKYYDRSDGLTNCLEYDQSNCLNEVK
 YDSDPALRAFLVEAGYIEAKFEAEENNTL"
 complement(4338..5693)
 /note="ORF_ID:s110996"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652230"
 /transl_table=11
 /translation="MNVSHRYHITIFGCGMKNKADSERMAGILENLTGMYTDDPNOAD
 LLYMTCSTRDAEOKVYSYLGROAKRKQREPELTLYVAGCAQOEBGLLRVPELD
 LVMQPHANRLDQLLEQVWAGSQVATBESLHIMEDITKRRSSTVSAMVNTITGCGER
 CSYCPVNRGVEQSRTPALIGEMEVLAAQGFKEVTLTGQNDIAYGRDLPPTSGSR
 BLHPTLDLYHYHDIEGIDRLRPAATSHRPTERLLOACQELPKYCEHFIPEQSGDN
 DLKAMKGYTREKYLQILEKIRRYMPDAIASDYVGFPEEAOFEENTLILIEVG
 FDLNTAAYSPPGTPPAFVWQNLSEYKQDRLRLNHLVSTQAMERSORLYRVEY
 LVEGENLKSPPQVMGRTKGNRLTFPGELSELGKTVPYKITEAPAFSLTGALSIVT
 A"
 5618..6969
 /gene="sqdB"
 /note="ORF_ID:s1r1020"
 /codon_start=1
 /product="sulfolipid biosynthesis protein sqdB"
 /db_xref="PID:g1652231"
 /transl_table=11
 /translation="MRAVYIGGDCGYGNATALYLKNKGYEVLIDSLVRYWDAOLGA
 ETLPIAPIRDLRWELTKIDLEIDINDYFELTNALROFOPDAVHVEGERSA
 PESMTDRHAVLTQANVNLNLALYALKEDPCHIVKLGTMGEYGNPIDIEGYSI
 TIEHGRKDTLPYKPGSGFYHLKSVHSHNHFKKIKGLATLNLNGIYGVUTTEE
 TGMDEMLNRDLDYDGVFGALNRFCIQAIHPLVFKKIGATLNLDRVLRITEL
 AIANPADKQGFVFNQYTELFEVGDLAQWQVAGADLGKVIDHLENPRVLEBHYF
 NAVNTNLDLGQIPRFLSDLSLNLNFKATYKXRDVQKHILPKYTWNG"
 7112..7402
 /note="ORF_ID:ssr1698"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652232"
 /transl_table=11
 /translation="MADPLTPAIDSRICKHMHMEDHASAIALYAYVGOOTDVTYMAQMO
 AIDPTGMDLVYESGSKTIRIEFEQPLKDSDAQVLIAMAKQARSVSKNS"
 complement(7439..8134)
 /note="ORF_ID:s110995"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652233"
 /transl_table=11
 /translation="MTVALDRRIITPDSGQPMADNTEQFENVILKENLECLFANHP
 DVFVGGDLMTVEBCHPETRVAPDVVALGRKGRKGRSROMQENNOAPQVFEILSP
 GNTLKEMTKLKEFDHGHVEEYVDPDNDELTGQRIGGETLIIEENAHVWSPLLGI
 KFEELSAETLRVYVDPGRFELSTVALATAEQASCRANEAOAEREKIRAEIAEAND
 RLKALLAAGIDV"
 complement(8220..9029)
 /note="ORF_ID:s110994"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652234"
 /transl_table=11
 /translation="MSTNFEFSTOHGVCOKTIVSRNMYIMFLHANTLYWLLG
 LGITLGLMITISGDEDDLDVDMKLEASSDVDTLHDELDDGDMGEAEVPMALQV
 LSFELGKVPMLILGIDFSLMGVILWNLVAVATGIMPELNGAGLIFLVSIAI
 SLWGLRSLRPAHLKFTFSQDVSTERYIGCGIVTSKRLPYLANGTIGQAHVYDNAG
 NLITISVSLPMAVYIPHNQBIILIIDSPKGYGLAIKAKSSDEDKWLKS"
 complement(8962..10059)

CDS
 /note="ORF_ID:s110993"
 /codon_start=1
 /product="potassium channel"
 /db_xref="PID:g1652235"
 /transl_table=11
 /translation="MGICSSQSENLNLIDROGRRLROELMAGATTLAGLFVYVGTAMY
 RYEDMTWMDAFYMTTITLATVGFGEFTHPLSPASLFTILIMLITIGYVWNPTE
 AFIOGYPDOSLRRROKOVIEBLADHYLLICGGRGGOOAFPEFAVENIDPVYDASPE
 VTIQKLRDIAVLOQDARLDELILLAHIHRAICISALSSDENITVLSATLWPKI
 RAIRAKSBEAVOKLKRAGADEVSPYITGSKRLAALRPQVVSFVGIITGADRSF
 YMEFRIGADEPYIGQTLREQLAQSALITAIARODRKLIVGPMGTHTLLDLSL
 ICLGVDELRLALNOLCLPILPAVRVLPKRNHR"
 10183..11472
 /gene="argD"
 /note="ORF_ID:s1r1022"
 /codon_start=1
 /product="N-acetylornithine aminotransferase"
 /db_xref="PID:g1652236"
 /transl_table=11
 /translation="MTVSPPVSVSAQAFVNDLSPAEFKADPDFTVMYNGREPI
 AIAKQSTLMDTEKSYLDYVAGIATCTTGHAFALYRAVSDQIKLHHVSNLYIP
 EGGELAKNIVESHSCADRVFCNSGAEANEAAIKVRYAHVYDLLEQPVITAKASF
 HGRTLAIITATGQPYQYQYFDPVYGFVYNDIRSLSENKRVADLEGNRYAAITFE
 PLOGSGVAPGDLAFKRVREICQNDLILVEDEVQVGRGKLMGEHGEEDIF
 TSAKGLAGVPIGAMACKKFCDFEPGNHASFPGNPLAACAGLAVLKTITGDRLLDN
 VQAREQORSGIAETIKNQXPILTFVPRWGLINGELISBESSIVSEIVKAMEGGL
 LAPAGPKVLRVYPLVTVTEAETQAVETLRQAIATLV"
 11599..11913
 /note="ORF_ID:s1r1023"
 /codon_start=1
 /product="hypothetical protein"
 /db_xref="PID:g1652237"
 /transl_table=11
 /translation="MKNIMQDRLPDQGFLLDLENNKTLERLITISGQITPPGEWY
 DQDREPVVLLQGRKAETIYDGAIVISLQSGDYLLIPAKQKHVFTTTPPCIMWLALH
 FS"
 11920..12561

CDS
 Note: remainder of annotations omitted.
 Query Match 1.6%; Score 24; DB 20; Length 150894;
 Best Local Similarity 81.6%; Pred. No. 1.38e+00;
 Matches 31; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

DB 4235 aaactctgatatggagccatgcatctatctcccca 4272
 OY 1388 AACCTATGATGTGGAGAAAGACATCACTCTCCCA 1425

RESULT 15
 LOCUS ECACEA 1344 bp DNA BCT 12-SEP-1993
 DEFINITION E. coli acea gene coding for isocitrate lyase (EC 4.1.3.1).
 ACCESSION X07543
 NID 940884
 KEYWORDS acea gene; isocitrate lyase; isocitritase.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Eubacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE
 AUTHORS Cozzone A.J.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-1988) Cozzone A.J., University of Lyon,
 Laboratory of Molecular Biology, 43 Boulevard du onze Novembre,
 69622 Villeurbanne, France
 2 (bases 1 to 1344)
 Riou C., Bleicher F., Duclos B., Cortay J.C. and Cozzone A.J.
 Nucleotide sequence of the acea gene coding for isocitrate lyase in
 Escherichia coli
 JOURNAL Nucleic Acids Res. 16 (12), 5689 (1988)
 MEDLINE 88262573
 FEATURES
 source Location/Qualifiers
 1..1344

RBS /organism="Escherichia coli"
 /strain="K12"
 /map="91 min"
 27..30
 /note="pot. RNA binding site"
 40..1344
CDS /note="isocitrate lyase (AA 1-434)"
 /codon_start=1
 /db_xref="PID:94085"
 /db_xref="SWISS-PROT:P05313"
 /translation="MKTRFQIQIEIQKEMTQPRWEGITRPYSAEDVYKLRGSVNPECT
 LAQLGAAMKRLHGESKRGYINSIGALTGQALQQAAGIEAVYLSGMQVADANMR
 PACIRISRSIRPOTCPAYERRINNTFRADQIOWSAGIEPGDPYVDYFLP IVADAEA
 GEGVINAFFELMKAMIEAGAAVHEEDOLASVKGGMGKVIPTQEAIOKIVPARL
 AADYGVPTLIVARTDADADLITSDCPYDSEPTIGERTSEGFRTTHAGIEQAISRG
 LAIAPYADLVNCESTIPDLELARRPAQAIHAKTFEKLAIYNSFSFNMOKNDDKTLA
 SFOQLSDMGYKFOFITLAGIHSWENFMDLANVAQEGMKHYEKVQPEPFAAKD
 GYTFVSHQOEVGTGFDFKVTIIQGTSSVTALIGSTEESQF"

COUNT 321 a 353 c 386 g 284 t

Query Match 1.5%; Score 23; DB 20; Length 1344;
Best Local Similarity 81.1%; Pred. No. 5.59e+00;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 983 cttataactgctgcgcgtcttcactgacagaaaaa 1019
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1488 CTTATTACTGCTTGCAGTATTCACTGGAAGAAAAA 1524

Search completed: Thu Aug 21 10:01:07 1997
Job time : 2597 secs.

THIS PAGE BLANK (USPTO)